International Journal of Mathematics and
Computer Applications Research (IJMCAR)
ISSN(P): 2249-6955; ISSN(E): 2249-8060
Vol. 6, Issue 6, Dec 2016, 1-10

APPLICATION OF ARIMA MODELS IN FORECASTING STOCK PRICES

SNEH SAINI[1], N. P. SINGH[2] & RATNA RAJ LAXMI[3]

[1]*Research Scholar, Maharishi Dayanand University, Rohtak, Haryana, India*

[2]*Professor, Management Development Institute, Gurgaon, Haryana, India*

[3]*Professor, Department of Statistics, Maharishi Dayanand University, Rohtak, Haryana, India*

ABSTRACT

For past three decade Autoregressive Integrated Moving Average Models are one of the most talked about models used for financial prediction in time –series analysis. This paper endeavors to present the forecasting of daily stock data of State Bank of India (SBI) for a period of 2 years. Box-Jenkins methodology of developing ARIMA models is used for forecasting the stock prices. The results reveals that ARIMA (0, 1, 1) is the most suitable model found for forecasting.

KEYWORDS: ARIMA Model, Stock Price, Forecasting Etc.

Received: Sep 21, 2016; **Accepted:** Oct 17, 2016; **Published:** Oct 20, 2016; **Paper Id.:** IJMCARDEC20161

Original Article

I. INTRODUCTION

Stock Prediction will always remain an interesting area for the researchers to explore. It has gathered the interest of the researchers to develop better predictive models every now and then. Stock price prediction is an important topic in finance and economics. The financial institutions and individual investors need an effective strategy to take decision everyday on the basis of predictions. The price prediction is regarded as one of most difficult task to achieve in financial forecasting due to complex nature of stock market [1, 2, 3] investment risk from the stock market. This remains an inspiring aspect for scholars to evolve new predictive models [4].

Till date lot many models and techniques had been advanced to predict stock prices. Among them ARIMA models are the most popular statistical tools. As it is quoted in the literature now a days forecasting can be done on the basis of two different viewpoints, one on the basis of Statistical Techniques, and other on the basis of Artificial Intelligence Technique [2]. ARIMA models are known to be very efficient for forecasting particularly for short-term prediction in the field of financial time series than even the most popular ANNs techniques [5, 6, 7].Other statistics models are regression method, exponential smoothing, generalized autoregressive conditional heteroscedasticity (GARCH). Few related works that has engaged ARIMA model for forecasting includes [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18].

This paper presents extensive process of building ARIMA models for short-term price prediction. The results obtained from real-life data established the potential strength of ARIMA models to provide stockholders short-term prediction that could support investment decision making process.

The rest of the paper is organized as follows. Section 2 presents brief overview of **Review of literature.** Section 3 describes ARIMA model. Section 4 presents the methodology used while section 5 illustrates the experimental results obtained. The paper is concluded in section 6.

II. REVIEW OF LITERATURE

Manoj and Edward J. (2016) applied ARIMAs models to forecast stock prices in their sector-specific study of forecasting and analyzing Indian sectoral stock prices. Their study suggested that ARIMA (1, 1, 0) is the most suitable model to be used for forecasting stock prices. It suggested that the stock prices are upward trends and could be considered as a worthy investment if examined closely. The ARIMA model was developed on the stock prices of 6 sectors viz. Banking, Healthcare, Automobiles, Power, Oil & Gas and IT. The developed models have one common characteristic i.e. they are all integrated at first order and are Autoregressive models with lag 1 having no MA characteristics. The data used in this study is of daily observations of past 9 years (Feb 2007 – April 2015) with 1996 observations.[11]

Adebiyi*et al.* (2014) carried out research on stock data obtained from New York Stock Exchange (NYSE) and Nigeria Stock Exchange (NSE). This paper presents extensive process of building stock price predictive model using the ARIMA model. The data used in the research are historical daily stock prices taken from stock exchanges of two countries from the period of 25th April, 1995 to 25th February, 2011 with 3990 observations. The results obtained from the paper using ARIMA model shows the potential of ARIMA models to predict stock prices on short-term basis satisfactorily.[14]

Kwasi and Kobina (2014) described modeling and forecasting of wholesale cassava monthly prices in central region of Ghana using ARIMA model. The experiment carried out on sample forecasting from January 2013 to December 2013. The study demonstrates a good performance in terms of explained variability and predicting power.[10]

Verma*et al.* (2015) analyzed ARIMA and state space modeling. These two procedures were compared for sugarcane yield modeling and forecasting to estimate the yield of sugarcane crop in three district of the Haryana, India. The time series yield data from 1960-61 to 2006-07 of sugarcane crop for the districts under consideration compiled from statistical abstracts of Haryana. The Study was primarily aimed at predicting a future value on the basis. Crop yield forecasting for the model testing period i.e. 2002-03,2003-04,2004-05,2005-06, 2006-07 and 2007-08 were achieved to check the validity of the models developed on the basis of sugarcane yield data from 1960-61 to 2001-02. ARIMA and State space models were found to be close to the real time yields and both the procedure were considered for the operational crop yield forecasting purpose. [8]

T. Ozaki (1977): Attempted to remove the difficulty in deciding the order of an ARIMA (autoregressive integrated moving average) model by using the MAICE (minimum AIC estimation) procedure, which selects a model by using Akaike's Information Criterion (AIC). The AIc, which is an estimate of the Kullback-Leibler information quantity, provides a powerful and almost automatic procedure for the identification of ARIMA model. It is remarkable that if we restrict ourselves, as was suggested by B-J, to the set of the five models (1, d, 0), (2, d, 0), (0, d, 1), (0, d, 2) and (1, d, 1), the MAICE procedure produces almost identical results as those obtained by the B-J procedure which requires highly skillful but nevertheless subjective judgment.[13]

Abdullahi, and Bakari (2014) examined the trend or pattern in the Nigeria capital market, as well as how to determine a suitable model for forecasting the Nigerian stock market by applying ARIMA model. Among the series of ARIMA models tested, it was discovered that ARIMA (2, 1, 2) model performs best since it has a minimum MAPE and MAE compared with the other models. The stock data used for this paper was taken from the period of 1985 to 2008. [12]

III. ARIMA MODEL

ARIMA model was first introduced by Box and Jenkins in 1970. It also referred to as Box-Jenkins methodology comprised of set of activities for identifying, estimating and diagnosing ARIMA models with time series data. The model is one of the most protuberant methods in financial forecasting [1, 6, 16]. ARIMA models have shown efficient competency to generate short-term forecasts. It constantly outperformed complex structural models in short-term prediction [17]. In ARIMA model, the future value of a variable is a linear combination of past values and past errors, expressed as follows:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \ldots - \theta_q \epsilon_{t-q}$$

where,

Y_t is the actual value, ϵ_t is the random error at t, Φ_i and θ_j are the coefficients, p and q are integers that are often referred to as autoregressive and moving average, respectively.

The steps in building ARIMA predictive model consist of: model identification, parameter estimation and diagnostic checking [18].

IV. METHODOLOGY

This research considered daily historical stock price data of State Bank of India (SBI) collected from Yahoo Finance covering the period starting from 20 January2014to 20 January 2016. In all, 523 observations were used for fitting various ARIMA models. On 31.10.2014 a split of shares of SBI was observed and data was adjusted accordingly to make the series uniform (*on web page this notification is reflected on 21st November 2014 but actual split on Yahoo Finance is considered on 31.10.14).* The data is comprised of four variables, namely: open price, low price, high price and close price respectively. In the research, open price of the data represents the price of the index to be forecasted.

The method used in the study to develop ARIMA model for stock price prediction is explained in detail underneath.

- The original data taken for the SBI stock is split in the ratio 1:10 for a period of 20.01.14 to 30.10.14.
- Data is tested for stationarity using Dickey Fuller Test.
- The selected data was found non- stationary then it was made stationary by taking first difference and second difference simultaneously.
- Then 17 ARIMA models were fitted taking p = [0,1,2], q = [0,1,2] and d= [0,1,2]
- Interpretation of statistics and analysis of graphs was done for identifying best ARIMA model.
- Inferences and conclusions were drawn accordingly.
- To determine the best ARIMA model among several experiments performed, the following criteria are used in this study for each stock index.

Figure 1 depicts the original pattern of the series to have general overview whether the time series is stationary or not. From the graph below the time series have random walk pattern.

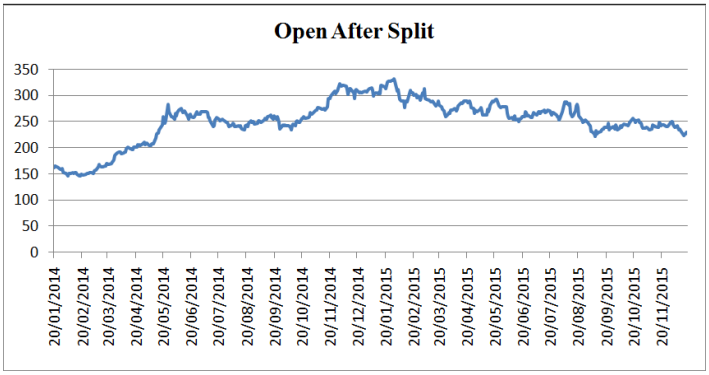


Figure 1: Opening Price of SBI Stock at BSE

It is evident from the graph above that the series is non-stationary. Therefore **Augmented Dickey-Fuller (ADF) unit root test** is applied on **opening price of SBI stock at BSE.** The result confirms that the series becomes stationary after the first- difference of the data. However the series is also checked at the second difference in anticipation of getting better results.

Table1: Dickey-Fuller Statistics

Dickey-Fuller Test			
	Original	**First Difference**	**Second Difference**
Tau (Observed value)	-1.500	-8.506	-14.756
Tau (Critical value)	-0.903	-0.907	-0.907
p-value (one-tailed)	0.826	< 0.0001	< 0.0001

Here **p-value (one-tailed) < alpha i.e.** < 0.0001 < 0.05. So the series is now stationary. It is very evident from graph below.

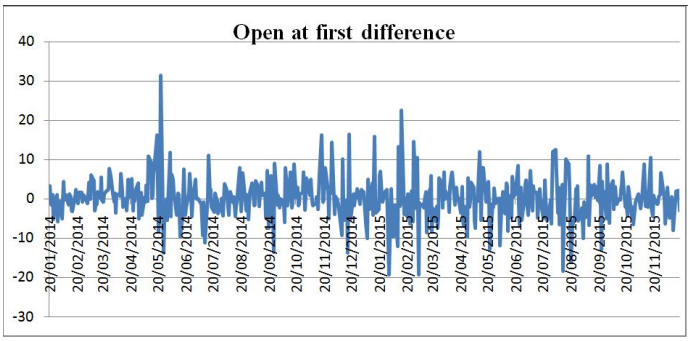


Figure 2: First Difference of Opening Price of SBI Stock at BSE

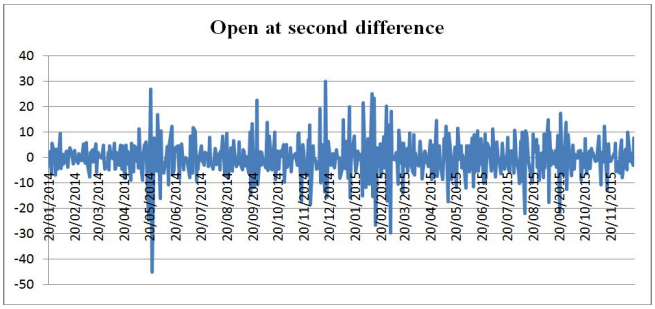


Figure 3: Second Difference of Opening Price of SBI Stock at BSE

Various ARIMA models are now applied to fit the best model. The results are shown in Table 2 below.

Table 2: Various Statistics of ARIMA Models on Original Values of Opening Price of SBI Stock at BSE

Model	(1,0,0)	(0,0,1)	(1,0,1)	(2,0,0)	(0,0,2)
SSE	14198.1	8E+06	14184	16382	2557886
MSE	28.3962	16872	28.368	32.764	5115.77
RMSE	5.32881	129.9	5.3261	5.724	71.5246
WN Variance	28.3962	16872	28.368	32.764	5115.77
MAPE(Diff)	1.46761	50.39	1.4674	1.6061	26.8252
MAPE	1.46761	50.39	1.4674	1.6061	26.8252
-2Log (Like.)	3099.4	6289	3098.9	3170.7	5695.11
FPE	28.5101	16872	28.481	33.027	5115.77
AIC	3103.4	6293	3104.9	3176.7	5701.11
AICC	3103.42	6293	3104.9	3176.8	5701.16
SBC	3111.82	6302	3117.5	3189.4	5713.75

Table3: Various Statistics of ARIMA Models for First & Second Difference of Opening Price of SBI Stock at BSE

Model	(0,1,0)	(1,1,0)	(0,1,1)	(2,1,0)	(0,1,2)	(1,1,1)	(0,2,0)	(1,2,0)	(0,2,1)	(1,2,1)	(2,2,1)	(1,2,2)
SSE	29225.2	21882	14211	19450	14193.9	14195	292	1E+05	87838	48721	37486.7	21962
MSE	58.5676	43.85	28.478	38.978	28.4448	28.447	587.7	253.7	176.38	97.833	75.2745	44.1
RMSE	7.65	6.62	5.33	6.24	5.33	5.33	24.24	15.93	13.28	9.89	8.67	6.64
WN Variance	58.56	43.85	28.47	38.97	28.44	28.44	587.77	253.7	176.38	97.83	75.27	44.1
MAPE (Diff)	4.4E+10	9E+11	5E+12	5E+11	4.7E+12	5E+12	99.8015	220	119.49	163.86	156.586	128.66
MAPE	598.76	642.5	242.13	687.18	249.28	248.5z	5.4E+12	1E+13	2E+12	8E+11	6.3E+12	1E+12
-2Log (Like.)	-	3303	3091.8	3244.5	3091.19	3091.2	-	4171	3995.5	3703.5	3574.31	3318.7
FPE	58.5676	44.03	28.478	39.292	28.4448	28.561	587.75	254.7	176.38	98.227	75.8816	44.278

AIC	-	3307	3095.8	3250.5	3097.19	3097.2	-	4175	3999.5	3709.5	3582.31	3326.7
AICC	-	3307	3095.8	3250.5	3097.24	3097.3	-	4175	3999.5	3709.6	3582.39	3326.8
SBC	-	3315	3104.2	3263.1	3109.83	3109.9	-	4184	4007.9	3722.2	3599.15	3343.5
Model	**(0,1,0)**	**(1,1,0)**	**(0,1,1)**	**(2,1,0)**	**(0,1,2)**	**(1,1,1)**	**(0,2,0)**	**(1,2,0)**	**(0,2,1)**	**(1,2,1)**	**(2,2,1)**	**(1,2,2)**

Interpretation

It can be inferred from the results obtained from table 3 that ARIMA (0, 1, 1) is the best fit model based on the values of FPE, AIC, AICC, SBC Tests. In all these tests the obtained values are minimal. It is quoted by T. Ozaki in his paper on the Order Determination of ARIMA Models (1977)

Table 4: Descriptive Statistics of Opening Price of SBI Stock at BSE

Model	Observations	Minimum	Maximum	Mean	Std. Deviation
(1,0,0)	500	147.00	332.80	252.71	41.87
(0,0,1)	500	147.00	332.80	252.71	41.87
(1,0,1)	500	147.00	332.80	252.71	41.87
(2,0,0)	500	147.00	332.80	252.71	41.87
(0,0,2)	500	147.00	332.80	252.71	41.87
(0,1,0)	500	-19.25	31.40	0.13	5.33
(1,1,0)	500	-19.25	31.40	0.13	5.33
(0,1,1)	500	-19.25	31.40	0.13	5.33
(2,1,0)	500	-19.25	31.40	0.13	5.33
(0,1,2)	500	-19.25	31.40	0.13	5.33
(1,1,1)	500	-19.25	31.40	0.13	5.33
(0,2,0)	500	-45.08	30.10	0.03	7.66
(1,2,0)	500	-45.08	30.10	0.03	7.66
(0,2,1)	500	-45.08	30.10	0.03	7.66
(1,2,1)	500	-45.08	30.10	0.03	7.66
(2,2,1)	500	-45.08	30.10	0.03	7.66
(1,2,2)	500	-45.08	30.10	0.03	7.66

Graphical Analysis if the Results Obtained

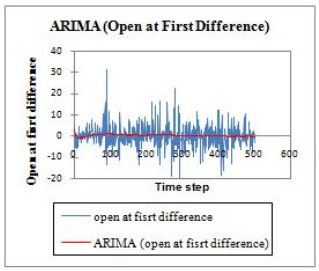
Moreover on observing the plot of the time series values over time to determine stationary or non-stationary, autocorrelation function (ACF) of the sample also reflects the data.

If the ACF of the time series values either cuts off or dies down fairly quickly, then values of the time series should be considered stationary.

Secondly if the ACF of the time series values either cuts off or dies down extremely slowly then it should be considered non-stationary.

Graphs below shows the series is stationary.

Results of ARIMA (0, 1, 1)



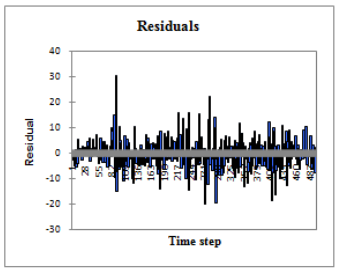
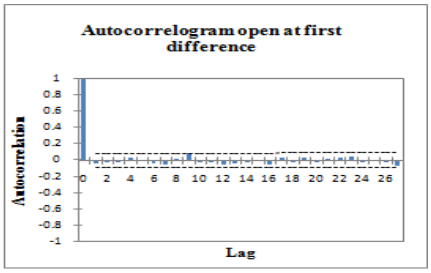


Figure 4: ARIMA-1D(0,1,1) **Figure 5: ARIMA-1D(0,1,1)**



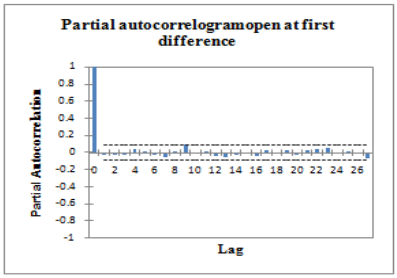
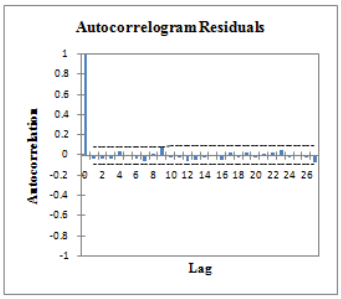


Figure 6: ARIMA-1D(0,1,1) **Figure 7: ARIMA-1D(0,1,1)**



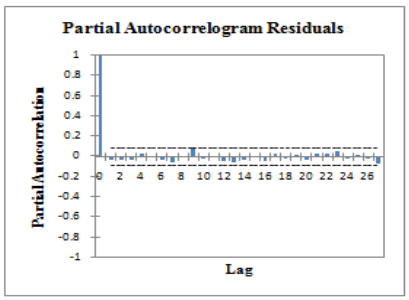


Figure 8: ARIMA-1D (0,1, 1) **Figure 9: ARIMA-1D(0,1,1)**

V. RESULTS AND DISCUSSIONS

The experimental results of SBI stock index are discussed in as below. Table 5 is the result of the predicted values of ARIMA (0, 1, 1) considered the best model for SBI stock index**.** Figure 10gives graphical illustration of the level accuracy of the predicted price against actual stock price to see the performance of the ARIMA model selected. From the graph it is obvious that the performance is satisfactory.

Table 5: Sample of Empirical Results of ARIMA (0, 1, 1) of SBI Stock index

Date	Actual Price (in Rs.)	Predicted Stock Price through ARIMA (in Rs.)
12/17/2015	230	224.54
12/18/2015	229.95	224.12
12/21/2015	226.9	223.15
12/22/2015	231.65	236.1
12/23/2015	230.25	225.34

Table 5: Contd.,		
12/24/2015	230.85	223.89
12/25/2015	228.3	227.12
12/28/2015	229.05	226.74
12/29/2015	228.85	234.23
12/30/2015	229	227.56
12/31/2015	225.95	228.12
1/1/2016	225	223.24
1/4/2016	226.95	225.12
1/5/2016	222.8	229.12
1/6/2016	217.9	215.45
1/7/2016	214	217.12
1/8/2016	212	214.54
1/11/2016	206.8	203.54
1/12/2016	204.85	200.87
1/13/2016	201.35	203.23
1/14/2016	197	198.96
1/15/2016	197.8	199.87
1/18/2016	184.9	186.4
1/19/2016	182.15	183.32
1/20/2016	180.65	178.12

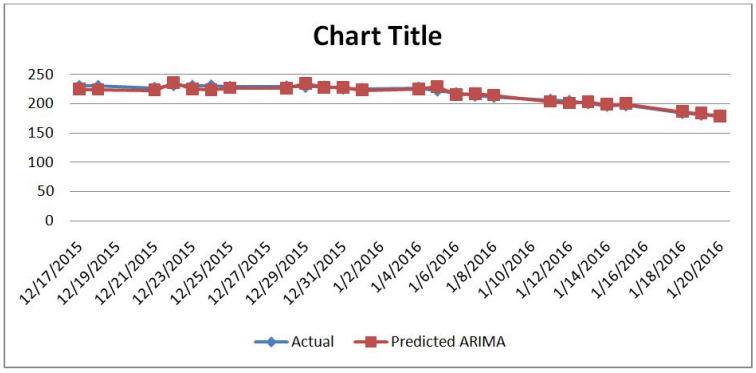


Figure 10: Graph of Actual Price vs. Predicted Values of SBI Stock Index

CONCLUSIONS

This paper presents extensive process of building ARIMA model for stock price prediction. The experimental results obtained from the daily data of State Bank of India (SBI) with best ARIMA model demonstrated the potential of ARIMA models to predict stock prices satisfactory on short-term basis. Box-Jenkins methodology of developing ARIMA models is used for forecasting the stock prices. ARIMA models have shown efficient capability to generate short-term forecasts. It constantly outperformed complex structural models in short-term prediction. This could guide investors in stock market to make profitable investment decisions. With the results obtained ARIMA models can compete reasonably well with emerging forecasting techniques in short-term prediction.

REFERENCES

1. *P. Pai and C. Lin, "A hybrid ARIMA and support vector machines model in stock price prediction", Omega vol.33 pp. 497-505, 2005*

2. *J.J. Wang, J.Z. Wang, Z.G. Zhang and S.P. Guo, "Stock index forecasting based on a hybrid model", Omega vol.40 pp.758-*

766, 2012.

3. *L.Y. Wei, "A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX", Economic Modelling vol. 33 pp. 893-899, 2013.*
4. *G.S. Atsalakis, E.M Dimitrakakis. and C.D. Zopounidis, "Elliot Wave Theory and neuro-fuzzy systems, stock market prediction: The WASP system", Expert Systems with Applications, vol. 38, pp.9196– 9206, 2011.*
5. *L.C. Kyungjoo, Y. Sehwan and J. John, "Neural Network Model vs. ARIMA Model In Forecasting Korean Stock Price Index (KOSPI),Issues in Information System, vol. 8 no. 2, pp. 372-378, 2007.*
6. *N. Merh, V.P. Saxena, and K.R. Pardasani, "A Comparison Between Hybrid Approaches of ANN and ARIMA For Indian Stock Trend Forecasting", Journal of Business Intelligence, vol. 3, no.2, pp. 23-43, 2010.*
7. *J. Sterba and Hilovska, "The Implementation of Hybrid ARIMA Neural Network Prediction Model for Aggregate Water Consumption Prediction", Aplim*
8. *U. Verma, A. Goyal and M. Goyal, "ARIMA versus State Space modelling: An application in agriculture", Indianjournals.com DOI: 10.5958/2349-2104.2015.00016.9 Adv. Appl. Res., Vol.7, No.2, (2015) pp 91 – 95*
9. *B. Uma Devi1 D.Sundar and Dr. P. Alli, "An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.3, No.1, January 2013*
10. *Bannor Richard Kwasi1 Bentil Julian Kobina, " Forcasting Of Cassava Prices In The Central Region Of Ghana Using Arima Model", Intercontinental Journal Of Marketing Research Review Volume 2, Issue 8 (August, 2014)*
11. *JyothiManoj and Aloysius Edward J., "ARIMA Modelling to Forecast and Analyze Indian Sectoral Stock Prices", International Journal in Management and Social Science (Impact Factor- 4.358), Vol.04 Issue-01 (January, 2016)*
12. *A.B, Abdullahi, H. R. Bakari, " Modelling the Nigerian Stock Market (Shares) Evidence from Time Series Analysis", The International Journal Of Engineering And Science (IJES) Volume – 3,Issue 4 Pages 01-12, 2014*
13. *T. Ozaki, "On the Order Determination of ARIMA Models", The Institute of Statistical Mathematics, Tokyo, Japan*
14. *Ayodele A. Adebiyi., Aderemi O. Adewumi, Charles K. Ayo, "Stock Price Prediction Using the ARIMA Model", 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation.*
15. *Theresa Hoang Diem Ngo, Warner Bros. Entertainment Group, Burbank, CA, " The Box-Jenkins Methodology for Time Series Models", Paper 454-2013*
16. *N. Rangan and N. Titida, "ARIMA Model for Forecasting Oil Palm Price", Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications, UniversitiSains Malaysia, 2006.*
17. *A.Meyler, G. Kenny and T. Quinn, "Forecasting Irish Inflation using ARIMA Models", Central Bank of Ireland Research Department, Technical Paper, 3/RT/1998.*
18. *B.G. Tabachnick and L.S. Fidell, "Using multivariate statistics", 4th ed., Person Education Company, USA 2001*